

REMARKS

Claim 1 has been amended and now recites that the first subset of historical data comprises less than all of the historical data; that the process for developing a predictive model is selected from among multiple possible model development processes; that the second subset of the historical data comprises less than all of the historical data; that the tentative predictive model is applied to the selected second subset; that a determination is made whether results of applying the tentative predictive model validate that the selected model development process will produce a final predictive model that is accurate for data that is not part of the historical data; and that, if the selected model development process is so validated, applying the validated model development process to the full set of historical data. None of these features was described in or would have been made obvious by the cited portions of the Lazarus reference.

Nor is this surprising, for Lazarus instead was interested in other features of model development. Accordingly, Lazarus described one process for developing a model, apparently used all of his historical data at once, and did nothing to validate his model development process by determining anything about the application of a tentative predictive model (developed using only part of the historical data) to another part of—less than all of—the historical data.

The additional comments of the applicant below are each preceded by related comments of the examiner (in small, bold type).

5. Claims 1-28 are rejected under 35 U.S.C. 102(b) as being anticipated by Lazarus et al. (U.S. Patent Number: 6,430,539). As to independent claim 1, Lazarus discloses a machine-base method comprising:

in connection with a project (e.g., predictive modeling of consumer financial behavior) (see Abstract), selecting a first subset of historical data (e.g., merchant vector) (see Figure 5, Col. 16, Lines 4-7, and Col. 37, Lines 54-65) for use in generating a predictive model (e.g. creates a predictive model) based on the historical data about a system being modeled (e.g., based on historical data) (see Col. 4, Lines 11 -16);

There is nothing in the figure 5 or text passages of Lazarus cited by the examiner that suggested that merchant vectors are less than all of the historical data, as now recited in the applicant's claim 1.

selecting a model development process from multiple model development processes (e.g., highest correlation with the segment vector, highest average transaction amount, or other selective criteria) (see Table 13; and Col. 37, Lines 54-65);

If any selection was described in the cited portions of Lazarus, it was a selection of a targeted population to which a finished model was to be applied, not a selection of a model development process, one that was later validated, as recited in the applicant's claim 1.

applying the selected model development process on the first subset of historical data (e.g., merchant vector) (see Figure 5, Col. 16, Lines 4-7, and Col. 37, Lines 54-65) to generate a tentative predictive model (e.g., segment models) (see Table 13 and Col. 38, Lines 23-38);

The cited passages did not describe and would not have made obvious a selection of a model development process, but rather a selection of a target audience to which to apply a model. And the merchant vectors were not a first subset of the historical data. Claim 1 recites validation of a tentative predictive model. Lazarus, by using all the historical data without a "holdout" subset did not validate a model development process, but rather showed accuracy and self-consistency of a model that was already developed.

validating the selected model development process (e.g., lift chart useful for validating the performance of the predictive models) (see Col. 34, Lines 20-23) based on a second subset of the historical data (e.g., merchant segments) (see Figure 5 and Col. 16, Lines 4-7), the second subset (e.g., merchant segments) being at least a portion of a complementary dataset of the first subset (e.g., merchant vectors) or being randomly selected from the historical data and independent of the first subset (see Figure 5 and Col. 16, Lines 4-7);

The cited passages did not describe and would not have made obvious a selection of a model development process, which would be validated based on a subset of the historical data, but rather simply a selection of a target audience to which to apply a model. Lazarus described his lift chart as being used to validate just the accuracy of the models, not the accuracy and generalizability of the model development process. Figure 5 and the cited passage in column 16 did not describe and would not have made obvious that the first set was a subset of the historical data. Nor did Lazarus describe or make obvious that the merchant segments were a subset of, or generated using a subset of, less than all of the historical data.

and applying the validated model development process to a full set of historical data that includes the first and second subsets (e.g., based on historical data) generate a final predictive model (e.g. creates a predictive model) (see Col. 4, Lines 11-16 and Figure 5).

The cited portions of column 4 of Lazarus did not describe and would not have made obvious that a validated model development process was, after being validated, applied to a full set of historical data. To the contrary, Lazarus described that the validated model (not a validated model development process) would be applied to actual spending, not historical data.

As to independent claim 6, Lazarus discloses a machine-based method comprising: in connection with a project (e.g., predictive modeling of consumer financial behavior) (see Abstract), selecting a model development process from multiple model development processes to apply on a subset of historical data to generate a first predictive model (e.g., highest correlation with the segment vector, highest average transaction amount, or other selective criteria) (see Table 13; and Col. 37, Lines 54-65), the selected model development process including automatic transformations of variables of the subset of the historical data (e.g., variables) (see Col. 11, Lines 13-23), automatic generation of the first predictive model (e.g., segment models) (see Table 13 and Col. 38, Lines 23-38), and automatic generation of performance measures of the first predictive model (e.g., confirm model performance) (see Col. 4, Lines 25-26), validating the selected development process based on the performance measures of the first predictive model (e.g., validation and analysis of the segment predictive models done to confirm model performance) (see Col. 11, Lines 21 -23); and

applying the validated model development process (e.g., lift chart useful for validating the performance of the predictive models) (see Col. 34, Lines 20-23; and Col. 38, Lines 23-38) to a full set of historical data (e.g., based on historical data) (see Col. 4, Lines 1-1 6) that includes the subset (see Figure 5 and Col. 16, Lines 4-7) to generate a second, final model (e.g. creates a predictive model) (see Col. 4, Lines 11 - 16).

The examiner has relied on the notion that Lazarus described (in the examiner's words, emphasis added) “validation and analysis of the segment predictive models done to confirm model performance”. These passages of Lazarus described validation of the models, not validation of a model development process, as recited in claim 6.

6. Applicant's amendment and arguments filed March 19, 2009 have been fully considered. The amendment does not overcome the original art rejection and the arguments are not persuasive. The following are the Examiner's observations in regard thereto.

Applicant Argues:

The examiner contended that Lazarus validated a model development process at "prior art Col. 4, Lines 25-26 where Lazarus discloses validation used to confirm model performance". However, confirming model performance does not mean validating the process of developing a model.

The model development process is not the generated model itself, but a process that generates the model. The model development process can include automatic transformations of variables of the data (e.g., specification, pages 11-16), automatic selection of transformed

variables (e.g., id.), automatic generation of a predictive model (e.g., specification, pages 22-23). Lazarus did not describe and would not have made obvious any consideration of whether the process of generating the predictive model was valid or not.

Examiner Responds:

Examiner is not persuaded. Using applicant's definition of the model development process to be "automatic generation of a predictive model (e.g., applicant's specification pages 22-23)," see Lazarus Figure 5 and Col. 16 Lines 4-7 for prior art anticipation of the predictive model generation system.

The examiner asserts that the applicant defined "the model development process" to be "automatic generation of a predictive model." The applicant strongly disagrees with this assertion. What the applicant referred to at pages 22 and 23 of its specification was "a rule-driven analytic workbench ... that assists analysts in development optimal predictive analytics models by automating ... model development process validation ..." The applicant is puzzled how the examiner could read this sentence to be a definition of "model development process" as "automatic generation of a predictive model."

Similarly, the examiner contends that figure 5 and the one sentence contained in the cited part of column 16 of Lazarus anticipated all of what is recited in the applicant's claim 1. Yet not a single element of claim 1, with the possible exception of the final clause, was described in or would have been obvious from that sentence of Lazarus.

The applicant can find nothing in figure 5 of Lazarus or the related cited text that described or would have suggested the use of less than all of the data in the master file to generate the merchant vectors or to generate the merchant segment predictive model. Nor can the applicant find anything in the cited portions of Lazarus to suggest that the merchant segment predictive model was a tentative model, as recited in the applicant's claim 1. The applicant's amended claim 1 recites selecting a first subset of historical data that is less than all of the historical data and a second independent subset that is less than all of the historical data.

Lazarus apparently used all the consumer transactions at all the merchants to define the merchant vectors that are transformed by clustering into merchant segments. Then Lazarus (Fig. 5) again used all the consumer transactions mapped onto the merchant segments to generate predictive consumer purchase segments. Lazarus apparently was forced to do so because his sole method for segmenting was clustering, which is severely impacted if cases are removed.

The applicant's claim 1 also recites that a process is selected for developing a predictive model for the system and that the process is selecting from among multiple possible model

development processes. In Lazarus there appears to be only one model generation process, not multiple ones, and no choice is made among multiple ones.

In claim 1, the selected model development process is applied to the first subset to generate the tentative predictive model. In Lazarus, there is only one model development process and the resulting model is not a tentative one.

Amended claim 1 refers to selecting a first subset of historical data about a system being modeled for use in generating a tentative predictive model for the system. Claim 1 recites that the subset comprises less than all of the historical data. Apparently, the examiner interprets figure 5 of Lazarus as showing the selection of a subset of the merchant data from the master file for use by the merchant vector generation module for 10 producing the merchant vectors for 16. In other words, the examiner apparently views the merchant vectors as a tentative predictive model, and the merchant data constituting a subset of historical data. However, it appears that Lazarus uses historical data about all of the merchants for which there is data in the master file in generating the merchant vectors.

In any case, the merchant vectors did not amount to a predictive model let alone a tentative predictive model. They were simply vectors that describe the merchants whose records are contained in the master file.

And according to independent claim 6 where "validating the selected development process" is "based on the performance measures of the predictive model," see Col. 34, Lines 20-23 for a lift chart useful for validating the performance of the predictive models. Under such considerations, the prior art anticipates the claims as written.

Again, the applicant respectfully believes that the examiner has incorrectly equated "validating the performance of the predictive models" with the applicant's validation of a model development process.

Applicant Argues:

In addition, Lazarus did not describe and would not have made obvious "selecting a model development process from multiple model development processes". Lazarus used one model generation module for generating his model (column 31, lines 30-35). Lazarus did not consider different development processes or select one of the multiple processes to generate his model.

Examiner Responds:

Examiner is not persuaded. See office action above for prior art anticipation of this newly presented claim limitation. See particularly Table 13 and Col. 37, Lines 54-65 for examples of model development processes, such as, "highest correlation with the segment vector, highest average transaction amount, or other selective criteria." Under such considerations, the prior art anticipates selecting a model development process from multiple model development processes.

What the examiner refers to as selecting a model development process is actually a description of how the models that Lazarus created could have been applied, once they were created, by a financial institution by identifying target populations.

Conclusion

7. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure.

Tang et al. (U.S. Publication Number: 200410002833) teaches systems and methods for mining model accuracy display for multiple state prediction.

8. Applicant's amendment necessitated the new ground(s) of rejection presented in this Office action. Accordingly, THIS ACTION IS MADE FINAL. See MPEP § 706.07(a). Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

Please apply \$555 for the Petition for Extension of Time fee and any other charges or credits to deposit account 06-1050, referencing attorney docket 17146-0008001.

Date: 01/11/10

Respectfully submitted,



David L. Feigenbaum
Reg. No. 30,378

Fish & Richardson P.C.
225 Franklin Street
Boston, MA 02110
Telephone: (617) 542-5070
Facsimile: (877) 769-7945
22338634.doc